



評価関数の複合化による神経回路網モデルの高性能化に関する研究

著者	早川 吉弘
号	21
発行年	2000
URL	http://hdl.handle.net/10097/13051

氏 名 (本 籍)	はやかわ 早 川	よしひろ 吉 弘	(宮城県)
学 位 の 種 類	博 士 (情報科学)		
学 位 記 番 号	情 第 21 号		
学位授与年月日	平成 13 年 3 月 13 日		
学位授与の要件	学位規則第 4 条第 2 項該当		
最 終 学 歴	平成 2 年 3 月東北大学大学院工学研究科博士課程前期 2 年の課程		
論 文 題 目	評価関数の複合化による神経回路網モデルの高性能化に関する研究		
査 委 員 (主 査)	東北大学教授 沢 田 康 次 東北大学教授 中 島 康 治		
	東北大学教授 星 宮 望 (工学研究科)		

論文内容要旨

生物は生存競争を勝ち抜くために、進化の過程で環境に非常に柔軟に適応できる情報処理機構である脳を獲得してきた。この我々の脳と現在のコンピューターとの最大の相違は、コンピューターは明確な目的を持って作られたものであり、その動作記述を明確に与えなければいけないのに対して、脳は進化の原動力である自然淘汰によって作られたものであり、常に環境に適応する必要性から自ら必要な機能を獲得・変化する情報処理装置である点にある。

神経回路網モデルの研究は、情報処理装置としての脳の最小単位である神経細胞を数値モデルで表現し、このモデル素子を用いて神経回路網を再構成する構成的手法により、現在のコンピューターとは全く異なる脳中の情報処理原理を解明・理解することであり、脳を真似た「脳型計算機」を実現することが目標とされる。この神経回路網モデルには並列分散処理による高速情報処理、自ら環境に適応して必要な機能を獲得出来る学習規則の存在、未知入力に対する汎化能力、シナプス荷重値への分散表現であるため故障に対するロバスト性を有するなどの、現在のコンピューターには無い魅力的な特質がある。

神経回路網モデルにおいては評価関数により、連想記憶のダイナミクスや学習ダイナミクス等が可能となる。しかし、評価関数の導入による神経回路網ダイナミクスにおける問題は極小値への捕捉と平坦領域の存在による情報処理過程の停滞等の性能劣化である。この二つの大きな問題を解決することが神経回路網モデルを高性能化するための重要な課題である。

本論文では、一般的解決手法として複数の評価関数の和を考慮することにより、極小値への捕捉、平坦領域の存在による情報処理過程の停滞の二つの問題を同時に解決することを

目的とする。具体的には、極小値への捕捉問題として、組み合わせ最適化問題への神経回路網モデルの適用問題を扱い、問題を記述する複数の評価関数の和の観点からこの問題の解決を行った。また、平坦領域の問題に対しては、学習則、特に応用範囲の広さから重要である Back-Propagation を扱い、新たな評価関数の付加によりこの問題の解決を行った。

本論文の内容は次の通り、6 章により構成されている。

第 1 章「緒言」では本研究の動機について述べた。

第 2 章「研究背景と目的」では、従来の研究について述べ、未解決な問題点を明らかにし、研究の目的について述べた。

第 3 章「神経回路網モデルによる組み合わせ最適化問題解法の高性能化」

本章ではまず組み合わせ最適化問題が、問題規模の増加とともに探索すべき解候補が指数関数的に増加するために、計算時間の爆発を招く非常に困難な問題であり、神経回路網モデルを使用すれば高速解探索が可能であることから大きな期待が寄せられているが、極小値の問題から、出力解が必ずしも正解では無いばかりでなく解として望ましくないものさえも多数出力されることを述べている。この極小値の問題に対して従来の研究では、確率的な揺らぎを導入する等の方法が提案されているが、高速性が犠牲にされてしまう問題点があった。

本章では、神経回路網モデルにおける組み合わせ最適化問題の解探索において、その高速解法を生かした一般的な設計指針を評価関数の複合化の観点で示した。

最初に神経回路網モデルに埋め込むことの出来る組み合わせ最適化問題を評価関数の複合化の観点から議論し、性質の異なるコスト条件を含まないクラス I の問題、コスト条件を含むクラス II の問題の 2 種類に分類を行った。

コスト条件を含まないクラス I の問題については、離散時間離散情報の神経回路網モデルを用いる事によって、正解のみの状態を安定に出来る評価関数の構成に関する次の 3 つの条件 (拘束条件の種類の制限、拘束条件の適用制限、評価関数複合化の係数設定条件) を明らかにし、その数学的証明を行った。また、具体的な適用例として N-Queen 問題と 4 色問題を神経回路網モデルに解かせ、正解のみを出力することを示した。

コスト条件を含むクラス II の問題については、連続時間連続情報モデルの神経回路網モデルを用いる事によって、解表現を満足しない状態 (エラー) が作られる理由の解明と統計的な手法によるエラーの出現確率の予測法を示した。また、評価関数複合化の係数と神経回路網モデルにより出力される解との関係を検討し、エラーの予測値から現実的な解を得るための評価関数の係数の設定方法を示した。特に巡回セールスマン問題を取り上げ、一般的に用いられる解表現を使う場合に問題の規模 (都市数) によらずに評価関数の複合化で用いる係数比を $B/A=0.54$ の値に設定する事が最適であることを明らかにした。

第 4 章「神経回路網の高速学習アルゴリズム」

本章ではまず応用範囲の広さから Back-Propagation は神経回路網モデルにとって重要な学習法であるが、学習の収束に膨大な時間が必要とされる問題が指摘されていること、BP の学習収束時間短縮の研究として、例えば学習係数の適応的变化規則による学習の高速化を目指した rprop や、ニュートン法の手法を導入して慣性項の係数を適応的に変更していく Quickprop が提案されているが、これらの方法は Batch 式の学習を仮定しており、On-Line 学習では適用が困難であることが述べられている。

しかしながら、実時間で環境に適応していくような学習を目指す場合には On-Line 式の学習が必要不可欠である。従って、On-Line 式の Back-Propagation の高速化手法の確立が未解決な問題である。

本章では、従来のBPで行われる与えられた学習データとの出力の誤差のみを最小にするばかりでなく、学習データから得られる目的関数の局所的な特徴の誤差も評価関数として導入するといった、評価関数の複合化により On-Line 式の学習の高速化を可能にする Local Feature Learning Algorithm について述べた。

次に開発した Local Feature Learning Algorithm の具体的な計算方法について述べ、荷重値更新時の計算において懸念される実計算時間の増加を見積もり、従来の Back-Propagation の荷重値更新に必要な実計算時間と同程度であることを数値実験により示した。

また、1 入力-1 出力の関数近似問題を用いて LFL の効果を数値実験により検証した。その結果、目的関数の局所的な特徴情報として傾き情報 (1 次微分情報) を導入するのみでも従来の Back-Propagation に対して約 30~60 倍の学習速度を実現した。しかも、Local Feature Learning Algorithm は特徴的な形状が増えるに従って、より高速化の効果が顕著に現れるといった特徴を確認する事が出来た。

最後に、Local Feature Learning Algorithm は評価関数を複合化することで新たな学習時の評価関数を構成するため、従来の高速学習アルゴリズムとの併用が可能であるといった特徴があり、数値実験によりこの効果を示した。

第5章「ポテンシャル系での学習ダイナミクスの考察による学習の高速化」

本章では、本研究で開発した Local Feature Learning Algorithm が評価関数の複合化により高速化を実現しているのでポテンシャル (評価関数) の複合化による学習ダイナミクスの変化に注目して様々な性質の議論を行った。

最初に Local Feature Learning Algorithm の従来の Back-Propagation には無い特異な学習過程についての現象を述べた。

次に、ポテンシャル系の重ね合わせの視点から、学習の初期過程において目的関数の局所的である微分情報の性質より、その評価関数 (ポテンシャル) の勾配が高次微分になるにつれてより急峻になることを示し、この影響によって学習初期の速度増加が起きることを明らかにした。

また、Local Feature Learning Algorithm は、目的関数の局所的な特徴を神経回路網モデルの出力と同じように適合させるための新たな評価関数を重ね合わせる (複合化) ことで新たな評価関数を構成するために、学習過程において従来の Back-Propagation と学習曲線が異なるものとなることを示した。

更に、学習時間の問題に対してその原因とされている学習の停滞を引き起こす平坦領域の問題であるが、Local Feature Learning Algorithm では評価関数を重ね合わせることで新たな評価関数 (ポテンシャル系) を構成すること、個々の局所の特徴に起因する評価関数でも平坦領域が存在するものと考えられるが、それぞれの平坦領域の生じる位置は異なることから、複合化された評価関数では平坦領域が小さくなり、その結果として従来の Back-Propagation よりも短時間でこの平坦領域を脱出できる性質を持つことを述べ、数値実験により確認した。

最後にポテンシャル系の学習ダイナミクスの考察から、Local Feature Learning Algorithm で最適な荷重値に早い時間で近づけ、その後通常の Back-Propagation に切り替えることで、より高速に高精度の関数近似を可能に出来ることを示した。

第6章「結言」では、本研究で得られた成果を総括した。

論文審査の結果の要旨

神経回路網モデルは並列分散処理による高速動作と汎化能力、環境に適応出来る学習能力を持つことから新しい計算原理として期待されている。一方、神経回路網モデルは荷重値空間や神経素子の空間における評価関数の極小値の存在と平坦領域による動作の停滞現象が、学習や連想記憶等の過程における障害となっていた。著者はこれらの問題の研究を、その一般的解決手法として複数評価関数の複合化の観点から進め、それを解決することに貢献した。本論文はこれらの成果をまとめたもので全編6章からなる。

第1章は緒言であり、本研究の動機について述べている。

第2章では、関連する研究分野の背景について述べ、未解決な問題を明らかにし、研究の目的について述べている。

第3章では、評価関数に生じる極小値への捕捉問題に対して最適化問題の解法を扱い、評価関数の複合化によって高速解法の特徴を生かした一般的な神経回路網モデルの設計指針を明らかにしている。コスト条件を含まない問題については、極小値を排除する評価関数の構成法を示し数学的な証明を行っている。また、コスト条件を含む問題については、不必要な極小値の生じる確率の予測法を考案し、実用的解を得るための評価関数の設定方法を示している。これらは神経回路による情報処理システムの設計指針を与える有用な成果である。

第4章では、評価関数に生じる平坦領域に起因する停滞現象の問題に対してBack-Propagationの学習速度問題を扱い、目的関数の高階微分を含む複合化評価関数を用いたLocal Feature Learning Algorithm (局所特徴学習アルゴリズム)を開発したことを述べている。局所的特徴情報として傾き情報のみの導入でも、従来のBack-Propagationに対して約60倍の学習速度を実現できること、学習すべき目的関数の形状変化が増えるに従ってこの効果が増大すること、及び、従来の高速学習アルゴリズムとの併用が可能であることを数値実験により示している。このアルゴリズムの開発は神経回路の高速学習の研究分野において有用な成果であると云える。

第5章では、Local Feature Learning Algorithmにおける評価関数の複合化から生じる学習ダイナミクスの変化について議論を行い、学習初期の速度増加の理由や学習過程を明らかにしている。また、本アルゴリズムは学習の停滞時間の短縮効果を持つこと、学習途中で評価関数を切り替えることにより従来のBack-Propagationに比較して学習誤差を4桁程度小さく出来ることを、数値実験により示している。

第6章は結言であり、本研究で得られた成果を総括している。

以上要するに本論文は、神経回路網モデルの情報処理において一般的な問題であった極小値への捕捉問題、動作の停滞問題を評価関数の複合化により解決することに貢献したものであり、システム情報科学の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。